WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report ⓘ

Oct 5, 2024 – 03:46 pm BST

PDB ID	:	1DX1
Title	:	BOVINE PRION PROTEIN RESIDUES 23-230
Authors	:	Lopez Garcia, F.; Zahn, R.; Riek, R.; Billeter, M.; Wuthrich, K.
Deposited on	:	1999-12-15

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
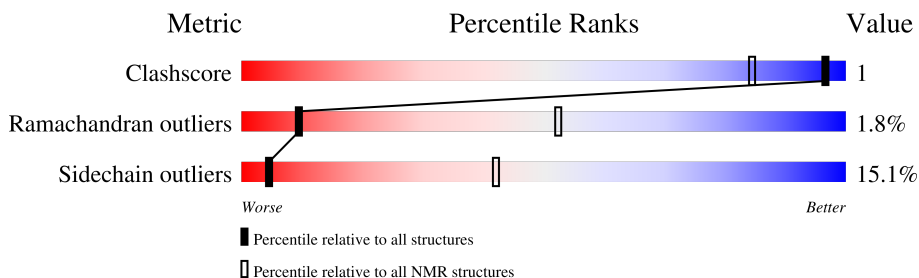http://www.wwpdb.org/validation/2017/FAQs#types.

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Clashscore		1
Ramachandran outliers		1.8%
Sidechain outliers		15.1%

Worse   *Better*

▮ Percentile relative to all structures

▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	219	42% • 53%

2　Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 18 is the overall representative, medoid model (most similar to other models). The authors have identified model 3 as representative.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:125-A:227 (103)	1.02	18

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters and 2 single-model clusters were found.

Cluster number	Models
1	2, 3, 5, 6, 8, 12, 13, 16, 17, 18
2	1, 9, 15, 19
3	10, 11
4	14, 20
Single-model clusters	4; 7

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 1675 atoms, of which 806 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PRION PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	104	1675	540	806	151	170	8	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	GLY	-	cloning artifact	UNP P10279
A	13	SER	-	cloning artifact	UNP P10279

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.
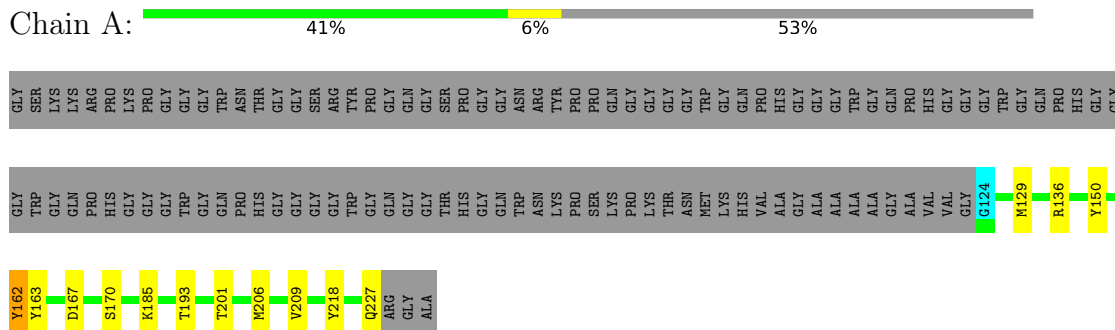
• Molecule 1: PRION PROTEIN
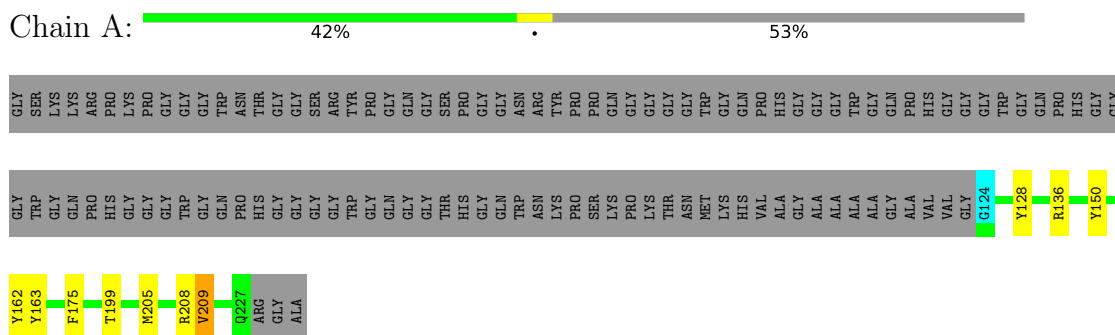
4.2 Scores per residue for each member of the ensemble
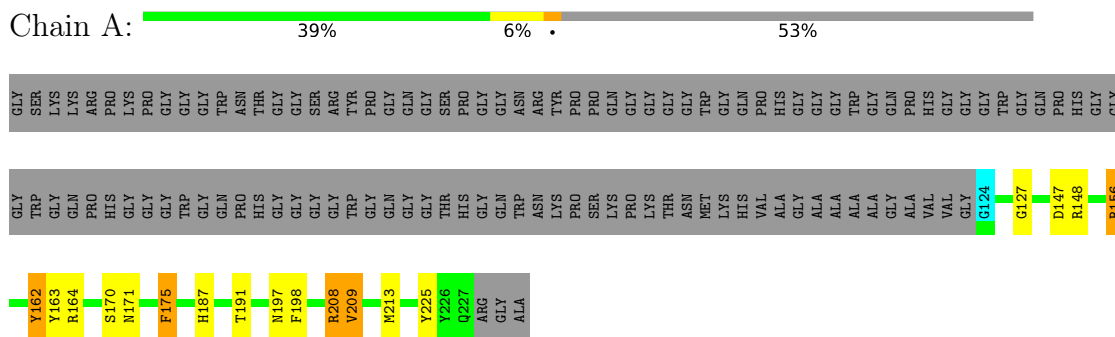
Colouring as in section 4.1 above.

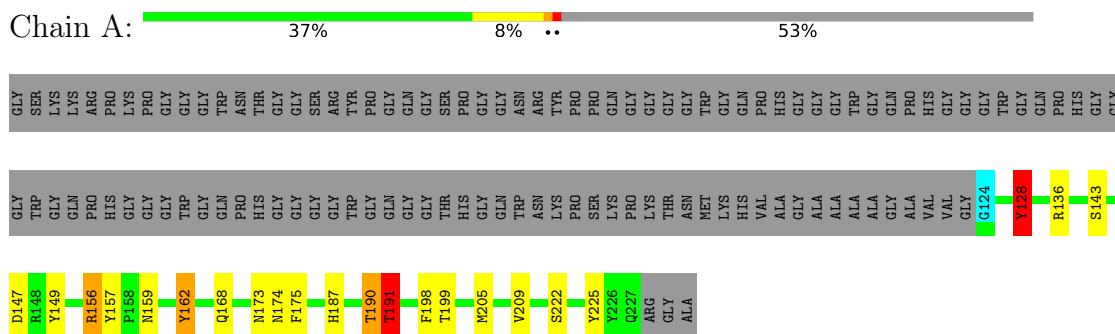### 4.2.1 Score per residue for model 1

• Molecule 1: PRION PROTEIN

4.2.2 Score per residue for model 2

● Molecule 1: PRION PROTEIN
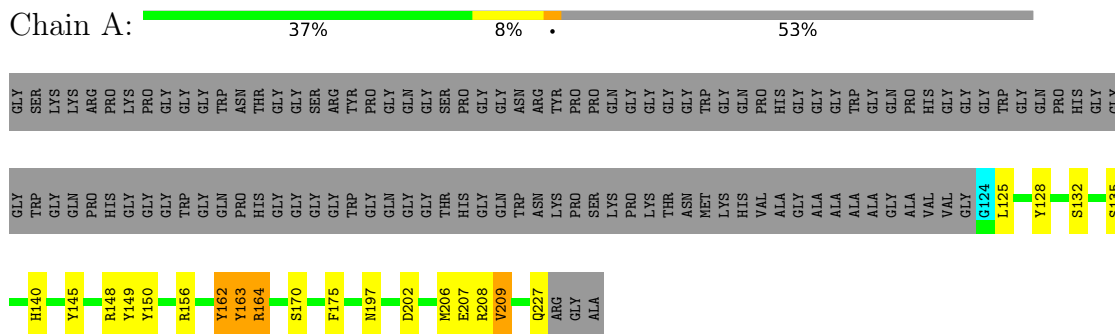
Chain A:

4.2.3 Score per residue for model 3

● Molecule 1: PRION PROTEIN

Chain A:

4.2.4 Score per residue for model 4

● Molecule 1: PRION PROTEIN

Chain A:

4.2.5 Score per residue for model 5

● Molecule 1: PRION PROTEIN

Chain A:	39%	6%	·	53%

4.2.6 Score per residue for model 6

• Molecule 1: PRION PROTEIN

Chain A:	36%	11%	53%

4.2.7 Score per residue for model 7

• Molecule 1: PRION PROTEIN

Chain A:	37%	10%	·	53%

4.2.8 Score per residue for model 8

• Molecule 1: PRION PROTEIN

Chain A:	37%	9%	53%

4.2.9 Score per residue for model 9

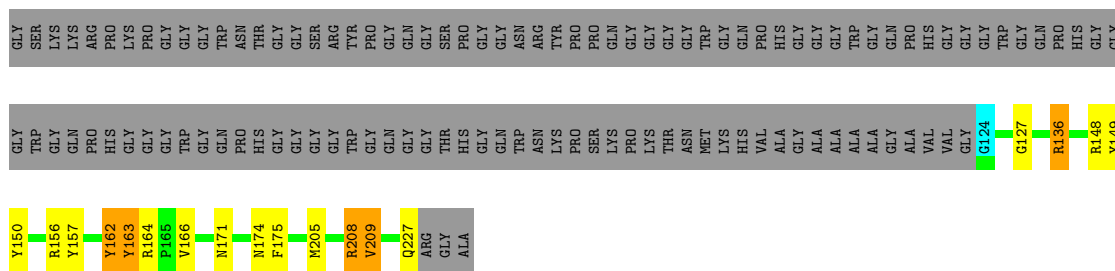• Molecule 1: PRION PROTEIN

Chain A: 38% 7% · 53%

4.2.10 Score per residue for model 10

• Molecule 1: PRION PROTEIN

Chain A: 37% 9% 53%

4.2.11 Score per residue for model 11

• Molecule 1: PRION PROTEIN

Chain A: 39% 7% · 53%

4.2.12 Score per residue for model 12

• Molecule 1: PRION PROTEIN

Chain A: 38% 7% · 53%

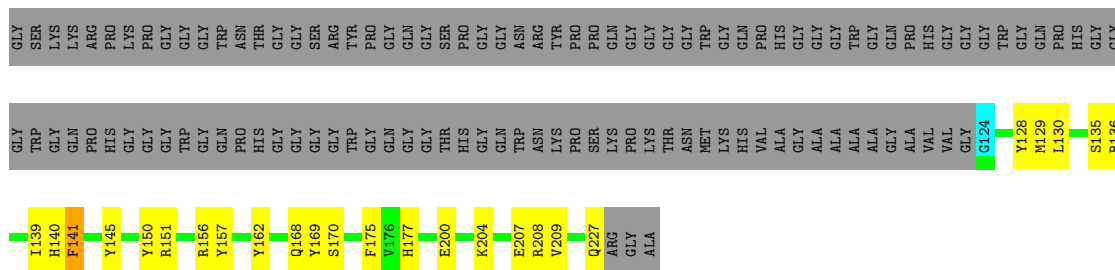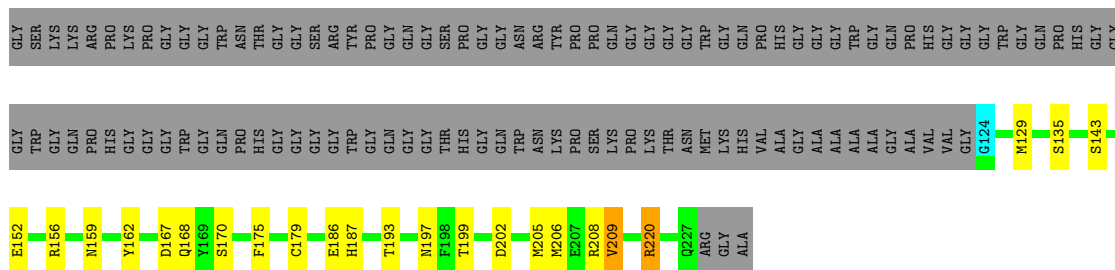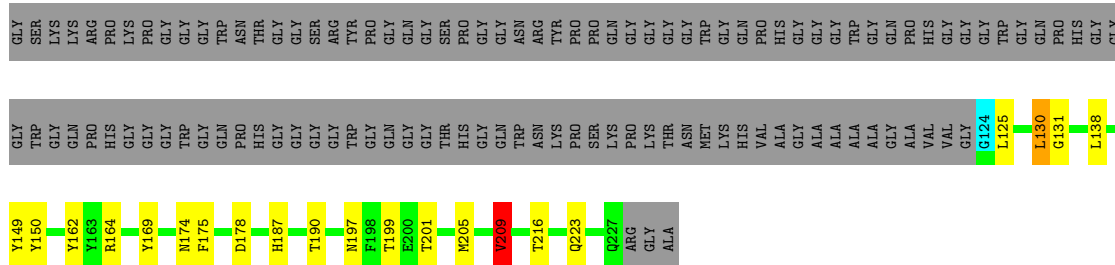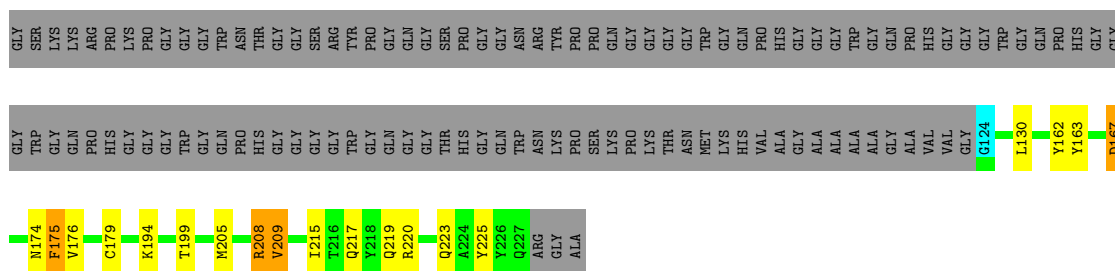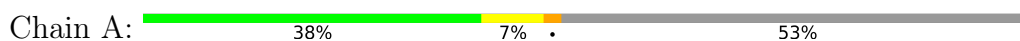#### 4.2.13 Score per residue for model 13

• Molecule 1: PRION PROTEIN

Chain A: 36% 9% · 53%
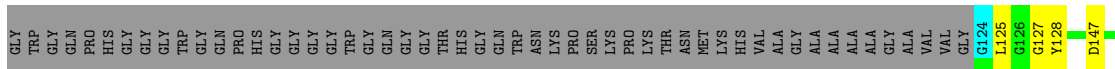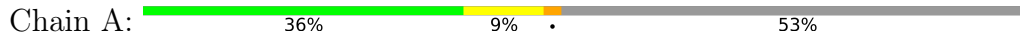
4.2.14 Score per residue for model 14

• Molecule 1: PRION PROTEIN

Chain A: 36% 11% 53%

M154 N159 Y162 Y163 R164 Q168 Y169 S170 F175 E186 T190 T191 T192 T193 K194 G195 E196 M205 M206 E207 R208 V209 Q217 Y218 Y226 Q227 ARG GLY ALA
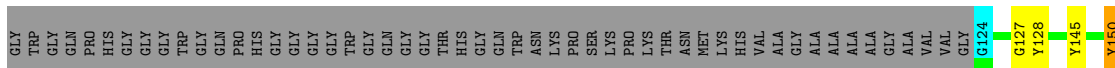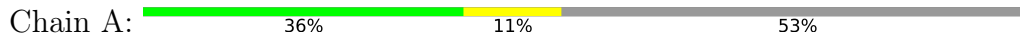
4.2.15 Score per residue for model 15

● Molecule 1: PRION PROTEIN

Chain A: 35%   12%   53%

GLY SER LYS LYS ARG PRO LYS PRO GLY GLY GLY TRP ASN THR GLY GLY SER ARG TYR PRO GLY GLN GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLY GLN GLY GLY GLY THR HIS GLY GLN TRP ASN LYS PRO SER LYS PRO LYS THR ASN MET LYS HIS VAL ALA GLY ALA ALA ALA GLY ALA VAL VAL GLY G124 Y128 M129 M134 S135 R136

P137 L138 I139 H140 Y145 Y149 Y150 Y157 Y162 Y163 R164 P165 V166 D167 S170 N171 F175 V184 T199 M205 M206 E207 R208 V209 C214 Q223 Q227 ARG GLY ALA

4.2.16 Score per residue for model 16

● Molecule 1: PRION PROTEIN

Chain A: 38%   8%   ·   53%

GLY SER LYS LYS ARG PRO LYS PRO GLY GLY GLY TRP ASN THR GLY GLY SER ARG TYR PRO GLY GLN GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLY GLN GLY GLY GLY THR HIS GLY GLN TRP ASN LYS PRO SER LYS PRO LYS THR ASN MET LYS HIS VAL ALA GLY ALA ALA ALA GLY ALA VAL VAL GLY G124 L125 Y128 M129 L130 G131 S132

A133 M134 S135 R136 P137 L138 D144 R151 V161 Y162 Q172 N173 N174 F175 V176 E186 M205 V209 I215 Y226 Q227 ARG GLY ALA

4.2.17 Score per residue for model 17

● Molecule 1: PRION PROTEIN

Chain A: 37%   8%   ·   53%

GLY SER LYS LYS ARG PRO LYS PRO GLY GLY GLY TRP ASN THR GLY GLY SER ARG TYR PRO GLY GLN GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLN PRO HIS GLY GLY GLY TRP GLY GLN GLY GLY GLY THR HIS GLY GLN TRP ASN LYS PRO SER LYS PRO LYS THR ASN MET LYS HIS VAL ALA GLY ALA ALA ALA GLY ALA VAL VAL GLY G124 L125 Y128 R136 P137 L138

I139 H140 R151 N159 Y162 Y163 R164 Q172 F175 I182 E186 T193 E200 R208 V209 I215 T216 Q217 Y218 Q219 Q227 ARG GLY ALA

4.2.18 Score per residue for model 18 (medoid)

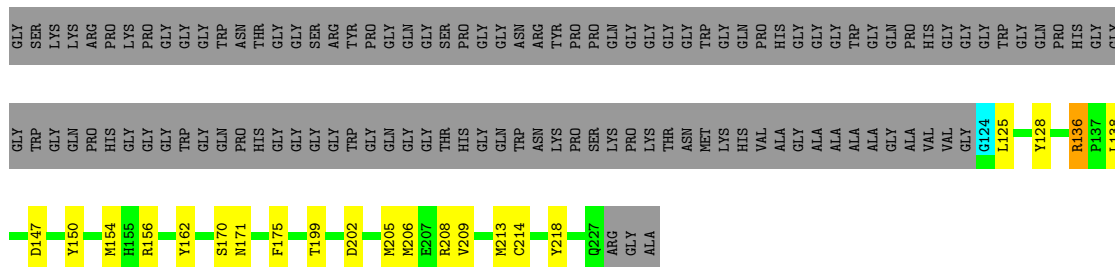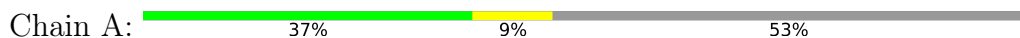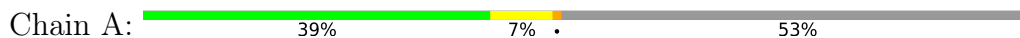• Molecule 1: PRION PROTEIN
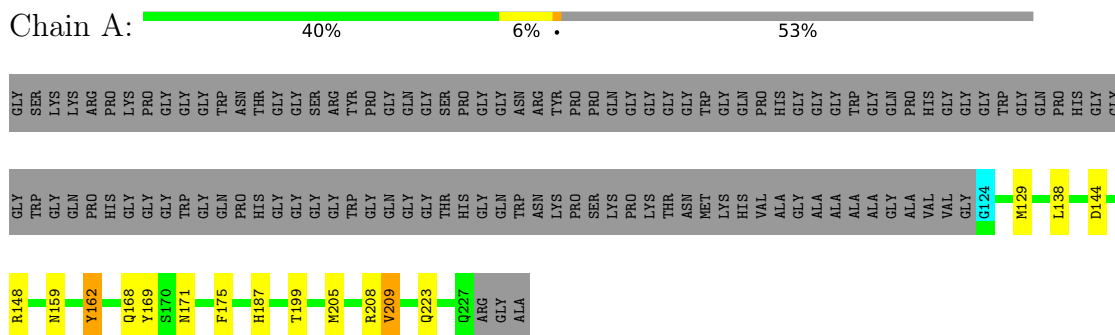
Chain A:

4.2.19 Score per residue for model 19

• Molecule 1: PRION PROTEIN
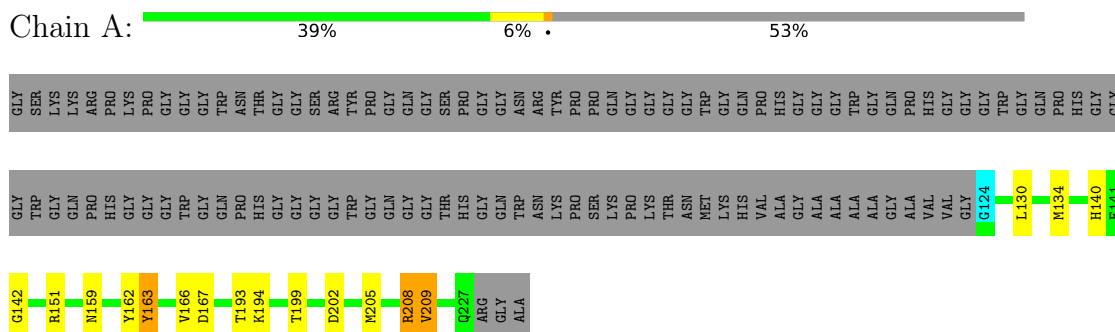
Chain A:

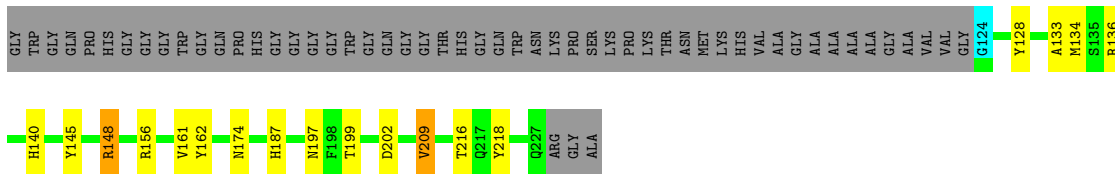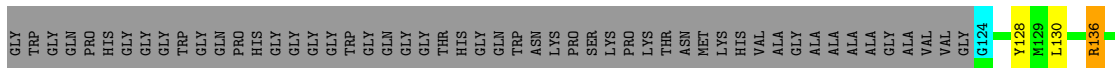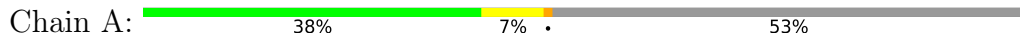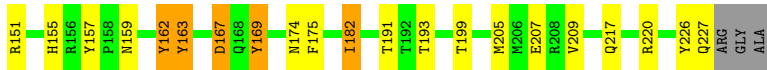### 4.2.20 Score per residue for model 20

• Molecule 1: PRION PROTEIN

Chain A:
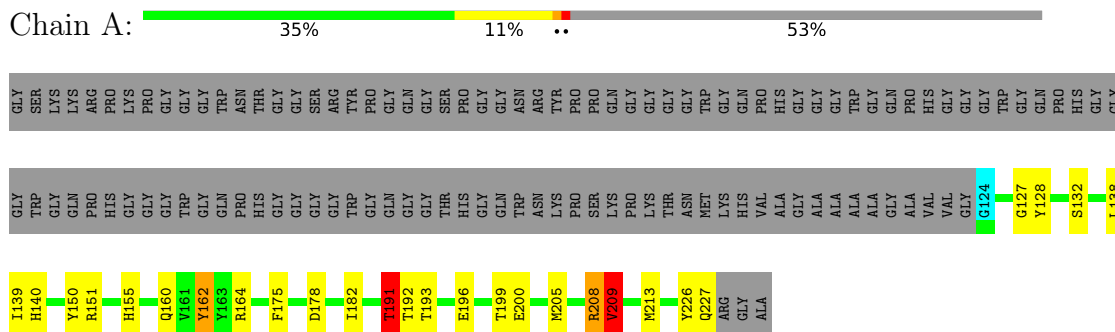
5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics.*

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION.*

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
OPALp	refinement	
DYANA	structure solution	

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.68±0.01	0±0/885 (0.0± 0.0%)	1.11±0.03	3±2/1196 (0.3± 0.2%)
All	All	0.68	0/17700 (0.0%)	1.11	61/23920 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	3.1±1.7
All	All	0	63

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
1	A	209	VAL	CA-CB-CG2	8.19	123.19	110.90	8	1
1	A	176	VAL	CA-CB-CG2	7.79	122.59	110.90	9	2
1	A	128	TYR	CB-CG-CD1	-7.23	116.66	121.00	3	2
1	A	162	TYR	CB-CG-CD2	-6.83	116.90	121.00	2	6
1	A	156	ARG	CD-NE-CZ	6.79	133.10	123.60	3	1
1	A	209	VAL	CA-CB-CG1	6.67	120.91	110.90	4	9
1	A	191	THR	CA-CB-OG1	6.60	122.86	109.00	20	2
1	A	164	ARG	NE-CZ-NH2	-6.40	117.10	120.30	8	2
1	A	175	PHE	CB-CG-CD1	-6.26	116.42	120.80	17	2
1	A	163	TYR	CB-CG-CD2	-6.13	117.32	121.00	5	4
1	A	220	ARG	NE-CZ-NH2	-6.12	117.24	120.30	7	1
1	A	156	ARG	NE-CZ-NH2	-6.04	117.28	120.30	2	2
1	A	151	ARG	NE-CZ-NH2	-5.97	117.31	120.30	17	2
1	A	208	ARG	NE-CZ-NH2	-5.90	117.35	120.30	12	4
1	A	157	TYR	CB-CG-CD2	-5.59	117.65	121.00	5	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Models Total
1	A	202	ASP	CB-CG-OD1	5.44	123.20	118.30	19	2
1	A	151	ARG	NE-CZ-NH1	5.44	123.02	120.30	19	2
1	A	169	TYR	CB-CG-CD2	-5.40	117.76	121.00	13	1
1	A	220	ARG	NE-CZ-NH1	5.39	123.00	120.30	9	1
1	A	208	ARG	NE-CZ-NH1	5.36	122.98	120.30	12	2
1	A	156	ARG	NE-CZ-NH1	5.30	122.95	120.30	4	2
1	A	148	ARG	NE-CZ-NH2	-5.25	117.68	120.30	11	2
1	A	148	ARG	NE-CZ-NH1	5.25	122.92	120.30	12	1
1	A	190	THR	CA-CB-CG2	-5.23	105.08	112.40	3	1
1	A	145	TYR	CB-CG-CD2	-5.20	117.88	121.00	4	1
1	A	136	ARG	NE-CZ-NH1	5.17	122.89	120.30	16	1
1	A	150	TYR	CB-CG-CD2	-5.07	117.96	121.00	1	1
1	A	175	PHE	CB-CG-CD2	-5.05	117.27	120.80	2	1
1	A	209	VAL	CG1-CB-CG2	-5.01	102.89	110.90	4	1
1	A	191	THR	CA-CB-CG2	5.00	119.41	112.40	3	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	150	TYR	Sidechain	9
1	A	218	TYR	Sidechain	5
1	A	156	ARG	Sidechain	4
1	A	208	ARG	Sidechain	4
1	A	128	TYR	Sidechain	4
1	A	149	TYR	Sidechain	4
1	A	157	TYR	Sidechain	4
1	A	136	ARG	Sidechain	4
1	A	145	TYR	Sidechain	4
1	A	151	ARG	Sidechain	4
1	A	162	TYR	Sidechain	3
1	A	164	ARG	Sidechain	3
1	A	163	TYR	Sidechain	3
1	A	225	TYR	Sidechain	2
1	A	220	ARG	Sidechain	2
1	A	148	ARG	Sidechain	2
1	A	171	ASN	Peptide	1
1	A	134	MET	Peptide	1

6.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	865	803	803	1±1
All	All	17300	16060	16060	24

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:206:MET:HA	1:A:209:VAL:HG22	0.69	1.64	14	2
1:A:201:THR:HG22	1:A:205:MET:SD	0.54	2.42	8	1
1:A:128:TYR:CE2	1:A:182:ILE:HG12	0.51	2.41	13	3
1:A:205:MET:O	1:A:209:VAL:HG22	0.50	2.07	8	1
1:A:206:MET:HA	1:A:209:VAL:CG2	0.50	2.36	10	2
1:A:172:GLN:O	1:A:176:VAL:HG13	0.50	2.06	16	1
1:A:191:THR:HG21	1:A:198:PHE:CE2	0.48	2.43	3	2
1:A:206:MET:HA	1:A:209:VAL:HG12	0.46	1.86	7	2
1:A:191:THR:HG23	1:A:192:THR:H	0.45	1.71	20	1
1:A:184:VAL:HG13	1:A:206:MET:SD	0.44	2.53	15	1
1:A:150:TYR:CE2	1:A:154:MET:CE	0.43	3.02	14	1
1:A:139:ILE:HD11	1:A:209:VAL:HA	0.43	1.91	20	1
1:A:209:VAL:HB	1:A:213:MET:SD	0.42	2.54	10	1
1:A:176:VAL:HG11	1:A:215:ILE:HG12	0.41	1.91	16	1
1:A:139:ILE:HB	1:A:141:PHE:CZ	0.41	2.50	6	1
1:A:215:ILE:HG22	1:A:219:GLN:HE21	0.41	1.74	9	2
1:A:130:LEU:HD22	1:A:131:GLY:N	0.40	2.32	8	1

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	102/219 (47%)	88±4 (86±4%)	12±3 (12±3%)	2±1 (2±1%)	9	52
All	All	2040/4380 (47%)	1762 (86%)	241 (12%)	37 (2%)	9	52

All 15 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	128	TYR	6
1	A	167	ASP	5
1	A	127	GLY	5
1	A	170	SER	4
1	A	166	VAL	3
1	A	171	ASN	3
1	A	168	GLN	2
1	A	125	LEU	2
1	A	141	PHE	1
1	A	133	ALA	1
1	A	192	THR	1
1	A	132	SER	1
1	A	172	GLN	1
1	A	142	GLY	1
1	A	191	THR	1

6.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	96/161 (60%)	82±3 (85±3%)	14±3 (15±3%)	4	42
All	All	1920/3220 (60%)	1631 (85%)	289 (15%)	4	42

All 66 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	162	TYR	20
1	A	175	PHE	17
1	A	209	VAL	17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	199	THR	12
1	A	205	MET	12
1	A	208	ARG	11
1	A	163	TYR	9
1	A	136	ARG	8
1	A	159	ASN	8
1	A	174	ASN	7
1	A	140	HIS	7
1	A	138	LEU	7
1	A	193	THR	6
1	A	227	GLN	6
1	A	187	HIS	6
1	A	197	ASN	6
1	A	130	LEU	6
1	A	129	MET	5
1	A	147	ASP	5
1	A	207	GLU	5
1	A	164	ARG	4
1	A	135	SER	4
1	A	170	SER	4
1	A	169	TYR	4
1	A	186	GLU	4
1	A	125	LEU	4
1	A	223	GLN	4
1	A	217	GLN	4
1	A	226	TYR	4
1	A	148	ARG	3
1	A	156	ARG	3
1	A	190	THR	3
1	A	191	THR	3
1	A	202	ASP	3
1	A	168	GLN	3
1	A	200	GLU	3
1	A	167	ASP	3
1	A	194	LYS	3
1	A	134	MET	3
1	A	182	ILE	3
1	A	213	MET	2
1	A	225	TYR	2
1	A	143	SER	2
1	A	132	SER	2
1	A	179	CYS	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	178	ASP	2
1	A	216	THR	2
1	A	171	ASN	2
1	A	214	CYS	2
1	A	161	VAL	2
1	A	160	GLN	2
1	A	155	HIS	2
1	A	196	GLU	2
1	A	144	ASP	2
1	A	185	LYS	1
1	A	201	THR	1
1	A	173	ASN	1
1	A	222	SER	1
1	A	149	TYR	1
1	A	206	MET	1
1	A	151	ARG	1
1	A	177	HIS	1
1	A	204	LYS	1
1	A	152	GLU	1
1	A	154	MET	1
1	A	172	GLN	1

6.3.3   RNA ⓘ

There are no RNA molecules in this entry.

6.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.5   Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.6   Ligand geometry ⓘ

There are no ligands in this entry.

6.7   Other polymers ⓘ

There are no such molecules in this entry.

6.8   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided